Sparsity-based Recovery of Finite Alphabet Solutions to Underdetermined Linear Systems

Abdeldjalil Aïssa-El-Bey *Senior Member, IEEE*, Dominique Pastor *Member, IEEE*, Si Mohamed Aziz Sbaï and Yasser Fadlallah *Member, IEEE*
Institut Télécom; Télécom Bretagne; UMR CNRS 6285 Lab-STIC, Technopôle Brest-Iroise CS 83818, 29238 Brest, France
Université européenne de Bretagne

***Abstract*—We consider the problem of estimating a deterministic finite alphabet vector $\boldsymbol{f}$ from underdetermined measurements $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{f}$, where $\boldsymbol{A}$ is a given (random) $n \times N$ matrix. Two new convex optimization methods are introduced for the recovery of finite alphabet signals via ℓ_1-norm minimization. The first method is based on regularization. In the second approach, the problem is formulated as the recovery of sparse signals after a suitable sparse transform. The regularization-based method is less complex than the transform-based one. When the alphabet size p equals 2 and (n, N) grows proportionally, the conditions under which the signal will be recovered with high probability are the same for the two methods. When $p > 2$, the behavior of the transform-based method is established. Experimental results support this theoretical result and show that the transform method outperforms the regularization-based one.**

I. Introduction

The source separation problem is an important research topic in a variety of fields, including speech and audio processing [1], radar processing [2], medical imaging [3], and communication [4]. As such, it has been intensively investigated in the literature in the past three decades. Basically, source separation aims to estimate original source signals from their mixtures. Approaches in this area can be classified according to the nature of the mixing process (instantaneous, convolutive) and the ratio between the number of sources and the number of sensors of the problem (determined, underdetermined, overdetermined). The more difficult case is clearly the underdetermined case, where the number of sources is more than the number of observed signals and for which solutions cannot be derived without additional assumptions. For instance, the sources can be separated thanks to their sparse representation in the time-frequency domain [5], [6], a source being said to be sparse in a given signal representation domain if most of its samples are close to zero. Another approach can be based on geometric properties of signals as in [7].

In the present paper, we address separation of finite alphabet signals in the instantaneous case with underdetermined known mixing matrix. This problem is important in data communications for symbol demodulation, in image processing and in operations research. In this respect, we pose the problem in the noiseless case as in [8]–[11]. However, in contrast to the aforementioned references that are tailored to alphabets with two elements only, we study the case of alphabets with any finite size without assuming that the signal to reconstruct is necessarily sparse. In this respect, we propose two criteria different from those introduced in [8]–[10]. Both criteria involve reformulating the initial problem so as to introduce sparsity constraints. We thus follow an approach similar to that proposed in [12] and [13], where an ℓ_0-norm optimization problem is relaxed into a sparse recovery problem involving ℓ_1-norm optimization. More specifically, we show that separating finite alphabet signals in the instantaneous case with underdetermined known mixing matrix can be rewritten as two distincts sparse recovery problems. Each of these problems can then be relaxed into ℓ_1-norm optimization. This convex relaxation provides good recovery performance for generic random mixing matrices [8], satisfying appropriate asymptotic properties.

This is mathematically proved for both methods, when the alphabet has two elements. For any alphabet size, the result is established for only one of the two methods, namely the transform-based method. Experimental results show that the transform-based method outperforms the other one, called the regularization-based method. The regularization-based method is however less computationally expensive than the transform-based one for any alphabet size above 2. When the alphabet size equals 2, the two methods have slightly the same complexity as that proposed in [8].

Henceforth, bold upper cases denote real-valued matrices. The transpose of a given matrix $\boldsymbol{A}$ is denoted by $\boldsymbol{A}^T$. All vectors will be column vectors unless transposed. Throughout the paper, $\mathbf{0}$ stands for the null vector and $\mathbf{1}_m$ is the (column) vector of $\mathbb{R}^m$ with one entries only. For a vector $\boldsymbol{x}$ the notation x_i will stand for the i^{th} component of $\boldsymbol{x}$. As usual, for any integer m, $[\![1, m]\!]$ stands for $\{1, 2, \dots m\}$.

II. Problem statement

The notation and terminology introduced in this section are used throughout the rest of the paper with always the same meaning. We hereafter consider the underdetermined linear system of equations or noise free mixing model

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}, \tag{1}$$

where $\boldsymbol{x} = [x_1, x_2, \cdots, x_N]^T$ is the $N \times 1$ source vector, $\boldsymbol{y} = [y_1, y_2, \cdots, y_n]^T$ is the $n \times 1$ observed vector and $\boldsymbol{A}$ is an $n \times N$ real-valued generic random matrix with $n < N$. For the sake of readiness, we recall the following definition.

Definition II.1 *: **Generic random matrices [8]** A given matrix $\boldsymbol{A}$ is an $n \times N$ generic random matrix if all sets of n columns are linearly independent with probability 1 and each column is symmetrically distributed about the origin.*

System (1) is underdetermined. To discard the case when it has no solution, we shall assume throughout that $\boldsymbol{A}$ has full-rank, which implies that the image of $\boldsymbol{A}$ is $\mathbb{R}^n$ and that the solutions of Eq. (1) form a vector space. In telecommunication systems, most signals are generated via some finite alphabet. For practical applications, it thus makes sense to assume that Eq. (1) has solutions in $\mathcal{F}^N$ where $\mathcal{F} = \{\alpha_1, \alpha_2, \cdots, \alpha_p\}$ is a given finite alphabet. We then study the conditions under which, given $\boldsymbol{y}$ such that:

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{f}, \quad \boldsymbol{f} \in \mathcal{F}^N, \tag{2}$$

$\boldsymbol{f}$ is actually the unique solution to this equation and how this unique solution can then be recovered in polynomial time under these conditions.

In this respect, we propose two new frameworks for the recovery of finite alphabet signals. These two frameworks are presented in Sections III. Both are based on reformulations of the finite alphabet constraints as sparsity constraints in incomplete measurements. The resulting sparse problems can then be relaxed as for sparse reconstructions in [12] and [13], by ℓ_1-norm optimization. Simulations results are given in Section V.

III. Sparsity-based recovery methods

A solution to Eq. (2) is given in [8] for the special case $\mathcal{F} = \{-1, 1\}$. This solution is obtained by solving the ℓ_∞-norm minimization

$$(P_\infty) \; : \qquad \underset{\boldsymbol{x} \in \mathbb{R}^N}{\arg\min} \|\boldsymbol{x}\|_\infty \quad \text{subject to} \quad \boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}, \tag{3}$$

since solutions in $\mathcal{F}^N$ to Eq. (2) are basically vertices of the hypercube $[-1, 1]^N$. Readily, (P_∞) can be solved by linear programming. A straightforward extension to any alphabet $\mathcal{F} = \{\alpha_1, \alpha_2\}$ with $\alpha_1 < \alpha_2$ is always possible via a simple translation. The solutions are then among the vertices of the hypercube $[\alpha_1, \alpha_2]^N$. Nevertheless, by construction, this method does not apply to alphabets with cardinality $p > 2$. Indeed, when $\mathcal{F} = \{\alpha_1, \alpha_2, \ldots, \alpha_p\}$ with $\alpha_1 < \alpha_2 < \ldots < \alpha_p$, the set of solutions in $\mathcal{F}^N$ to Eq. (2) may involve any point of the hypercube $[\alpha_1, \alpha_p]^N$.

To overcome this limitation, we hereafter propose two criteria aimed at solving Eq. (2) by seeking sparse solutions of a possibly transformed linear system of equations. The first method presented in the next section is regularization-based. The second one in Section III-B involves a suitable sparse transform.

A. Regularization-based method

Our first approach is to consider the finite alphabet constraint as prior knowledge to incorporate into the penalty function. A basic way to involve this constraint is simply to use the ℓ_0-norm to count the coordinates of any given $\boldsymbol{x} \in \mathbb{R}^N$ that do not belong to $\mathcal{F}$. Thence, the use of $\sum_{i=1}^{p} \|\boldsymbol{x} - \alpha_i \mathbf{1}_N\|_0$ as the cost to minimize. Actually, we have the following result.

Proposition III.1 *Given $\boldsymbol{f} \in \mathcal{F}^N$ with $\mathcal{F} = \{\alpha_1, \alpha_2, \ldots, \alpha_p\}$, $\boldsymbol{f}$ is the unique solution in $\mathcal{F}^N$ to Eq. (2) if and only if $\boldsymbol{f}$ is the unique solution to the optimization problem:*

$$(P_{\mathcal{F},0}) : \quad \underset{\boldsymbol{x} \in \mathbb{R}^N}{\arg\min} \sum_{i=1}^{p} \|\boldsymbol{x} - \alpha_i \mathbf{1}_N\|_0 \;\; \textit{subject to} \;\; \boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}.$$

Proof *Let $\boldsymbol{x} \in \mathbb{R}^N$ such that $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$. For any $i \in [\![1, p]\!]$, let*

$$N_i = \operatorname{card}\{j \in [\![1, N]\!] : x_j = \alpha_i\}.$$

Then,

$$\sum_{i=1}^{p} \|\boldsymbol{x} - \alpha_i \mathbf{1}_N\|_0 = Np - \sum_{i=1}^{p} N_i \geq N(p-1).$$

Equality is attained when $\sum_{i=1}^{p} N_i = N$ and, thus, when $\boldsymbol{x} \in \mathcal{F}^N$. It follows that the solutions in $\mathcal{F}^N$ to Eq. (2) are exactly the solutions in $\mathbb{R}^N$ to $(P_{\mathcal{F},0})$, which straightforwardly leads to the conclusion.

Solving a ℓ_0-norm minimization problem is generally complex and may require exhaustive search strategy, which can be intractable in practice for large values of N or p. Therefore, by mimicking literature on sparse reconstruction [14], we propose to replace the ℓ_0-norm by the ℓ_1-norm. Thus, we address the much simpler problem:

$$(P_{\mathcal{F},1}) : \quad \underset{\boldsymbol{x} \in \mathbb{R}^N}{\arg\min} \sum_{i=1}^{p} \|\boldsymbol{x} - \alpha_i \mathbf{1}_N\|_1 \;\; \text{subject to} \;\; \boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}.$$

On the practical side, unlike the ℓ_0-norm, the ℓ_1-norm is convex. Furthermore, $(P_{\mathcal{F},1})$ can be solved by linear programming and, thus, in polynomial time. However, problem $(P_{\mathcal{F},1})$ does not always have the same solution as $(P_{\mathcal{F},0})$. For the special case $p = 2$, the next result leads to the conditions under which $(P_{\mathcal{F},1})$ yields the unique solution to Eq. (2) with high probability. To state this result, we need the notion of proportional growth, of which we hereafter recall the definition for readiness.

Definition III.1 *: **Proportional Growth Setting [15]** A sequence of couples (n, N_n) will be said to grow proportionally if there is $\delta \in (0, 1)$ so that $\frac{n}{N_n} \to \delta$ when $n \to +\infty$. To alleviate the notation, subscript n will henceforth be omitted.*

Theorem III.1 *For $\mathcal{F} = \{\alpha_1, \alpha_2\}$ ($p = 2$) where $\alpha_1 < \alpha_2$, we have:*

- ***(i)*** *$\boldsymbol{f} \in \mathcal{F}^N$ is the unique solution to Eq. (2) if and only if $\boldsymbol{f}$ is the unique solution to $(P_{\mathcal{F},1})$;*
- ***(ii)*** *If $\boldsymbol{A}$ is an $n \times N$ generic random matrix, the probability that $(P_{\mathcal{F},1})$ has a unique solution in $\mathcal{F}^N$ is $P_{n,N}$ given by*

$$P_{n,N} = 2^{-N+1} \sum_{i=0}^{n-1} \binom{N-1}{i}. \tag{4}$$

When (n, N) grows proportionally, this probability tends to 0 when $\frac{n}{N} < \frac{1}{2}$ and tends to 1 when $\frac{n}{N} > \frac{1}{2}$.

Proof *By using the triangular inequality, we have that*

$$\forall \boldsymbol{x} \in \mathbb{R}^N, \quad \|\boldsymbol{x} - \alpha_1 \mathbf{1}_N\|_1 + \|\boldsymbol{x} - \alpha_2 \mathbf{1}_N\|_1 \geq N(\alpha_2 - \alpha_1)$$

with equality if and only if, for every $i \in [\![1, N]\!]$,

$$|x_i - \alpha_1| + |x_i - \alpha_2| = \alpha_2 - \alpha_1.$$

Since $[\alpha_1, \alpha_2] = \{t \in \mathbb{R} : \ |t - \alpha_1| + |t - \alpha_2| = \alpha_2 - \alpha_1\}$, it follows that:

$$\arg\min_{\boldsymbol{x} \in \mathbb{R}^N} \Big(\|\boldsymbol{x} - \alpha_1 \mathbf{1}_N\|_1 + \|\boldsymbol{x} - \alpha_2 \mathbf{1}_N\|_1 \Big) = [\alpha_1, \alpha_2]^N$$

Therefore, when $p = 2$, solving $(P_{\mathcal{F},1})$ is equivalent to solving

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}, \quad \textit{subject to} \quad \boldsymbol{x} \in [\alpha_1, \alpha_2]^N. \tag{5}$$

By setting $\boldsymbol{x}' = \frac{(\boldsymbol{x} - \alpha_1 \mathbf{1}_N) + (\boldsymbol{x} - \alpha_2 \mathbf{1}_N)}{\alpha_2 - \alpha_1}$, the linear problem (5) is equivalent to:

$$\boldsymbol{y}' = \boldsymbol{A}\boldsymbol{x}', \quad \textit{subject to} \quad \boldsymbol{x}' \in [-1, 1]^N, \tag{6}$$

*with $\boldsymbol{y}' = \frac{1}{\alpha_2 - \alpha_1} (2\,\boldsymbol{y} - (\alpha_2 + \alpha_1)\boldsymbol{A}\,\mathbf{1}_N)$. Statement **(i)** follows.*

*According to [8], the probability that there exists a unique solution in $\{-1, 1\}^N$ to Eq. (6) is $P_{n,N}$ given by Eq. (4). Thence, statement **(ii)**.*

The behavior of the regularization-based method for $p > 2$ is an open issue. However, the experimental results provided in Section V show that this method performs less well than the transform-based method proposed in the next section.

B. Transform-based method

In this section, we propose a solution based on a suitable sparse transform so as to benefit, as in [16], from the combination of sparsity and finite alphabet constraints. To do so, we plunge $\mathcal{F}^N$ into $\mathbb{R}^{Np}$, so that any element $\boldsymbol{f} \in \mathcal{F}^N$ is represented by a sparse vector $\boldsymbol{s}(\boldsymbol{f}) \in \mathbb{R}^{Np}$. This sparse vector is composed of N consecutive p-uples, such that each p-uple contains one 1 only and $p - 1$ zeros. By so proceeding, the initial problem (2) becomes equivalent to a problem of sparse signal recovery from highly incomplete measurements.

To do so, for any given $\boldsymbol{f} \in \mathcal{F}^N$, we define:

$$\boldsymbol{s}(\boldsymbol{f}) = [\boldsymbol{\epsilon}_1, \boldsymbol{\epsilon}_2, \ldots, \boldsymbol{\epsilon}_N]^T \in \mathbb{R}^{Np} \tag{7}$$

with $\boldsymbol{\epsilon}_i = [\mathrm{I}(f_i = \alpha_1), \mathrm{I}(f_i = \alpha_2), \cdots, \mathrm{I}(f_i = \alpha_p)]$ and where $\mathrm{I}(f_i = \alpha_j)$ is the indicator function equal to one if $f_i = \alpha_j$ and zero otherwise. This transform is a complete disjunctive coding such as that used in data analysis.

We also introduce the matrices $\boldsymbol{B}_\alpha$ and $\boldsymbol{B}_1$ in $\mathbb{R}^{N \times Np}$ defined by:

$$\boldsymbol{B}_\alpha = \boldsymbol{I}_N \otimes \boldsymbol{\alpha}^T \quad \text{and} \quad \boldsymbol{B}_1 = \boldsymbol{I}_N \otimes \mathbf{1}_p^T, \tag{8}$$

where $\boldsymbol{\alpha} = [\alpha_1, \cdots, \alpha_p]^T$ and $\otimes$ is the Kronecker product. By construction, we have:

$$\boldsymbol{f} = \boldsymbol{B}_\alpha \boldsymbol{s}(\boldsymbol{f}) \quad \text{and} \quad \boldsymbol{B}_1 \boldsymbol{s}(\boldsymbol{f}) = \mathbf{1}_N. \tag{9}$$

Therefore, we have:

$$\boldsymbol{\Phi}\, \boldsymbol{s}(\boldsymbol{f}) = \boldsymbol{b} \text{ with } \boldsymbol{\Phi} = \begin{pmatrix} \boldsymbol{A}\boldsymbol{B}_\alpha \\ \boldsymbol{B}_1 \end{pmatrix} \text{ and } \boldsymbol{b} = \begin{pmatrix} \boldsymbol{y} \\ \mathbf{1}_N \end{pmatrix}. \tag{10}$$

For any $i \in [\![1, N]\!]$, we henceforth put $T_i = [\![(i-1)p+1, ip]\!]$. Note that, for any $\boldsymbol{u} = [u_1, u_2, \ldots, u_{Np}] \in \mathbb{R}^{Np}$, $\sum_{j \in T_i} u_j$ is the i^{th} coordinate of $\boldsymbol{B}_1 \boldsymbol{u}$. This simple remark will prove helpful in the sequel.

On the basis of the above transform, the following proposition states that the recovery of finite alphabet signal $\boldsymbol{f}$ amounts to recovering the sparse signal $\boldsymbol{s}(\boldsymbol{f})$ from measurements $\boldsymbol{b}$.

Proposition III.2 *If $\boldsymbol{f} \in \mathcal{F}^N$ is the unique solution to Eq. (2), vector $\boldsymbol{s}(\boldsymbol{f})$ given by definition (7) is the unique solution to the optimization problem:*

$$(TP_{\mathcal{F},0}): \quad \arg\min_{\boldsymbol{u} \in \mathbb{R}^{Np}} \|\boldsymbol{u}\|_0 \ \textit{ subject to } \ \boldsymbol{\Phi}\boldsymbol{u} = \boldsymbol{b}.$$

Proof *: Let $\boldsymbol{u} \in \mathbb{R}^{Np}$ be any solution to $(TP_{\mathcal{F},0})$. We thus have $\boldsymbol{B}_1 \boldsymbol{u} = \mathbf{1}_N$, which implies that $\|\boldsymbol{u}\|_0 \geq N$. With the notation of the statement, it follows from Eqs. (7) and (10) that $\boldsymbol{s}(\boldsymbol{f})$ is solution to $(TP_{\mathcal{F},0})$ since $\|\boldsymbol{s}(\boldsymbol{f})\|_0 = N$. Therefore, $\|\boldsymbol{u}\|_0 = N$.*

Now, set $\boldsymbol{u} = [u_1, u_2, \ldots, u_{Np}]$. Since $\boldsymbol{B}_1 \boldsymbol{u} = \mathbf{1}_N$ and $\|\boldsymbol{u}\|_0 = N$, all the values u_j for $j \in T_i$ are null except one, which equals 1. Therefore, $\boldsymbol{B}_\alpha \boldsymbol{u} \in \mathcal{F}^N$. Moreover, because $\boldsymbol{u}$ is solution to the optimization problem $(TP_{\mathcal{F},0})$, $\boldsymbol{B}_\alpha \boldsymbol{u}$ satisfies Eq. (2). Since $\boldsymbol{f}$ is assumed to be the unique solution of (2), it follows that $\boldsymbol{B}_\alpha \boldsymbol{u} = \boldsymbol{f}$. According to the first equality in (9), we thus have $\boldsymbol{B}_\alpha \boldsymbol{u} = \boldsymbol{B}_\alpha \boldsymbol{s}(\boldsymbol{f})$. Taking into account that $\|\boldsymbol{s}(\boldsymbol{f})\|_0 = \|\boldsymbol{u}\|_0 = N$, we conclude that $\boldsymbol{u} = \boldsymbol{s}(\boldsymbol{f})$.

When Eq. (2) has a unique solution $\boldsymbol{f}$ in $\mathcal{F}^N$, the recovery of $\boldsymbol{f}$ may require an exhaustive search strategy to seek the unique solution to $(TP_{\mathcal{F},0})$ before applying the linear transform $\boldsymbol{B}_\alpha$ to this solution. For the same reasons as those evoked in the previous section, we aim at reducing the complexity cost of the optimization by relaxing the ℓ_0-norm into the ℓ_1-norm. In this respect, consider the optimization problem:

$$(TP_{\mathcal{F},1}): \quad \arg\min_{\boldsymbol{u} \in \mathbb{R}^{Np}} \|\boldsymbol{u}\|_1 \ \text{ subject to } \ \boldsymbol{\Phi}\boldsymbol{u} = \boldsymbol{b}.$$

Although the set of solutions to $(TP_{\mathcal{F},1})$ is not guaranteed to involve $\boldsymbol{s}(\boldsymbol{f})$ only, it is however expected that $\boldsymbol{f} = \boldsymbol{B}_\alpha \boldsymbol{s}$ for any solution $\boldsymbol{s}$ to $(TP_{\mathcal{F},1})$. Theorem III.2 answers to this point. In particular, it gives a necessary and sufficient condition under which a unique solution to Eq. (2) is the unique element of the image by $\boldsymbol{B}_\alpha$ of the set of all the solutions to $(TP_{\mathcal{F},1})$.

Theorem III.2 *With the notation introduced above, let $\mathcal{S}$ stand for the set of solutions in $\mathbb{R}^{Np}$ to $(TP_{\mathcal{F},1})$. Suppose that $\boldsymbol{f} \in \mathcal{F}^N$ is a solution to Eq. (2). Then:*

(i) *Vector $\boldsymbol{s}(\boldsymbol{f})$ defined by Eq. (7) belongs to $\mathcal{S}$,*

(ii) $\boldsymbol{B}_\alpha(\mathcal{S}) = \boldsymbol{f} + \boldsymbol{B}_\alpha \left(\mathrm{Ker}\, \boldsymbol{\Phi} \cap \mathcal{K}_{\boldsymbol{f}}^{N,p} \right)$, *where*

$$\mathcal{K}_{\boldsymbol{f}}^{N,p} = \big\{ \boldsymbol{h} \in \mathbb{R}^{Np} : \forall i \in [\![1, N]\!], -1 \leq h_{n_i(\boldsymbol{f})} \leq 0 \ \& \ \forall j \in T_i \setminus \{n_i(\boldsymbol{f})\}, h_j \geq 0 \ \& \ \sum_{j \in T_i} h_j = 0 \big\} \tag{11}$$

and, for every given $i \in [\![1, N]\!]$, $n_i(\boldsymbol{f})$ is the unique element of T_i such that the $n_i(\boldsymbol{f})^{th}$ coordinate of $\boldsymbol{s}(\boldsymbol{f})$ is 1.

(iii) *$\boldsymbol{f}$ is the unique element of $\boldsymbol{B}_\alpha(\mathcal{S})$ if and only if*

$$\operatorname{Ker} \boldsymbol{\Phi} \cap \mathcal{K}_{\boldsymbol{f}}^{N,p} \subset \operatorname{Ker} \boldsymbol{B}_\alpha. \tag{12}$$

Proof *Let $\boldsymbol{u} \in \mathbb{R}^{Np}$. By the triangle inequality, we have*

$$\|\boldsymbol{u}\|_1 = \sum_{i=1}^{Np} |u_i| = \sum_{i=1}^{N} \sum_{j \in T_i} |u_i| \geq \sum_{i=1}^{N} \left| \sum_{j \in T_i} u_i \right|.$$

Since $\sum_{j \in T_i} u_i$ is the i^{th} coordinate of $\boldsymbol{B}_1 \boldsymbol{u}$, $\sum_{i=1}^{N} \left| \sum_{j \in T_i} u_i \right| = \|\boldsymbol{B}_1 \boldsymbol{u}\|_1$. Therefore:

$$\|\boldsymbol{u}\|_1 \geq \|\boldsymbol{B}_1 \boldsymbol{u}\|_1 . \tag{13}$$

For all $\boldsymbol{u} \in \mathbb{R}^{Np}$ such that $\boldsymbol{\Phi u} = \boldsymbol{b}$, we have $\boldsymbol{B}_1 \boldsymbol{u} = \boldsymbol{1}_N$ and thus, according to (13)

$$\boldsymbol{\Phi u} = \boldsymbol{b} \quad \Rightarrow \quad \|\boldsymbol{u}\|_1 \geq N. \tag{14}$$

Suppose that $\boldsymbol{f} \in \mathcal{F}^N$ is a solution to Eq. (2). By construction of $\boldsymbol{s}(\boldsymbol{f})$, $\|\boldsymbol{s}(\boldsymbol{f})\|_1 = N$. It thus follows from (10) and (14) that $\boldsymbol{s}(\boldsymbol{f})$ is a solution to $(TP_{\mathcal{F},1})$. Therefore, any solution $\boldsymbol{s}$ to $(TP_{\mathcal{F},1})$ satisfies $\|\boldsymbol{s}\|_1 = N$. According to the foregoing that the set $\mathcal{S}$ of solutions to $(TP_{\mathcal{F},1})$ can be written as

$$\begin{aligned} \mathcal{S} &= \left\{ \boldsymbol{s}(\boldsymbol{f}) + \boldsymbol{h} : \ \boldsymbol{\Phi h} = \boldsymbol{0} \ \text{ and } \ \|\boldsymbol{s}(\boldsymbol{f}) + \boldsymbol{h}\|_1 = N \right\} \\ &= \boldsymbol{s}(\boldsymbol{f}) + \left(\operatorname{Ker} \boldsymbol{\Phi} \cap \mathcal{C}_{\boldsymbol{f}}^{N,p} \right), \end{aligned} \tag{15}$$

with $\mathcal{C}_{\boldsymbol{f}}^{N,p} = \left\{ \boldsymbol{h} \in \mathbb{R}^{Np} : \ \|\boldsymbol{s}(\boldsymbol{f}) + \boldsymbol{h}\|_1 = N \right\}$ and the convention $\boldsymbol{x} + \mathcal{D} = \{\boldsymbol{x} + \boldsymbol{d} : \ \boldsymbol{d} \in \mathcal{D}\}$ for any $\boldsymbol{x} \in \mathbb{R}^{Np}$ and any $\mathcal{D} \subset \mathbb{R}^{Np}$. According to appendix A, we have

$$\operatorname{Ker} \boldsymbol{\Phi} \cap \mathcal{C}_{\boldsymbol{f}}^{N,p} = \operatorname{Ker} \boldsymbol{\Phi} \cap \mathcal{K}_{\boldsymbol{f}}^{N,p}. \tag{16}$$

From (15) and (16), we derive that

$$\mathcal{S} = \boldsymbol{s}(\boldsymbol{f}) + \left(\operatorname{Ker} \boldsymbol{\Phi} \cap \mathcal{K}_{\boldsymbol{f}}^{N,p} \right).$$

Thereby, according to Eq. (9),

$$\begin{aligned} \boldsymbol{B}_\alpha(\mathcal{S}) &= \boldsymbol{B}_\alpha \, \boldsymbol{s}(\boldsymbol{f}) + \boldsymbol{B}_\alpha \left(\operatorname{Ker} \boldsymbol{\Phi} \cap \mathcal{K}_{\boldsymbol{f}}^{N,p} \right) \\ &= \boldsymbol{f} + \boldsymbol{B}_\alpha \left(\operatorname{Ker} \boldsymbol{\Phi} \cap \mathcal{K}_{\boldsymbol{f}}^{N,p} \right) \end{aligned}$$

Then, $\boldsymbol{f}$ is the unique element of $\boldsymbol{B}_\alpha(\mathcal{S})$ if and only if $\boldsymbol{B}_\alpha \left(\operatorname{Ker} \boldsymbol{\Phi} \cap \mathcal{K}_{\boldsymbol{f}}^{N,p} \right) = \{\boldsymbol{0}\}$, which is equivalent to $\operatorname{Ker} \boldsymbol{\Phi} \cap \mathcal{K}_{\boldsymbol{f}}^{N,p} \subset \operatorname{Ker} \boldsymbol{B}_\alpha$.

When $p = 2$, the necessary and sufficient condition stated by Theorem III.2 reduces to the equality $\mathcal{K}_{\boldsymbol{f}}^{N,2} \cap \operatorname{Ker}(\boldsymbol{A}\boldsymbol{B}_\alpha) = \{\boldsymbol{0}\}$, so that $\mathcal{S}$ contains $\boldsymbol{s}(\boldsymbol{f})$ only. This is established in Appendix B.

Similarly to Theorem III.1, the next result now gives the conditions under which, when Eq. (2) has a unique solution, this unique solution can be derived with high probability from the solutions to $(TP_{\mathcal{F},1})$.

Theorem III.3 *With the notation of Theorem III.2 and given $\mathcal{F} = \{\alpha_1, \alpha_2, \ldots, \alpha_p\}$ with $p \geqslant 2$, we have:*

(i) *$\boldsymbol{f} \in \mathcal{F}^N$ is the unique solution to Eq. (2) if and only if $\boldsymbol{B}_\alpha(\mathcal{S}) \cap \mathcal{F}^N = \{\boldsymbol{f}\}$,*

(ii) *If $\boldsymbol{A}$ is an $n \times N$ generic random matrix and Eq. (2) holds with $\boldsymbol{f}$ randomly chosen with uniform distribution in $\mathcal{F}^N$, the probability that $\boldsymbol{B}_\alpha(\mathcal{S}) \cap \mathcal{F}^N = \{\boldsymbol{f}\}$ is given by:*

$$Q_{n,N}(p) = \sum_{k=0}^{n-1} \binom{N}{k} \left(\frac{2}{p}\right)^{N-k} \left(\frac{p-2}{p}\right)^{k} (1 - P_{N-n,N-k}) \tag{17}$$

where $P_{n,N}$ is given by Eq. (4),

(iii) *Under the hypotheses of statement **(ii)** and by assuming that (n, N) grows proportionally, $Q_{n,N}(p)$ tends to 0 when $\frac{n}{N} < \frac{p-1}{p}$ and tends to 1 when $\frac{n}{N} > \frac{p-1}{p}$.*

Proof

***Proof of statement* (i):** *We begin with the direct implication. Suppose that $\boldsymbol{f}$ is the unique element of $\mathcal{F}^N$ that satisfies Eq. (2). According to Theorem III.2 (ii), $\boldsymbol{f}$ is an element of $\boldsymbol{B}_\alpha(\mathcal{S}) \cap \mathcal{F}^N$ since the null vector is an element of $\mathcal{K}_{\boldsymbol{f}}^{N,p}$. If $\boldsymbol{f}' \in \boldsymbol{B}_\alpha(\mathcal{S}) \cap \mathcal{F}^N$, Theorem III.2 (ii) also implies the existence of $\boldsymbol{h} \in \operatorname{Ker} \boldsymbol{\Phi} \cap \mathcal{K}_{\boldsymbol{f}}^{N,p}$ such that $\boldsymbol{f}' = \boldsymbol{f} + \boldsymbol{B}_\alpha \boldsymbol{h}$. Since $\boldsymbol{h} \in \operatorname{Ker} \boldsymbol{\Phi}$, $\boldsymbol{A}\boldsymbol{B}_\alpha \boldsymbol{h} = \boldsymbol{0}$. It follows that $\boldsymbol{f}'$ satisfies Eq. (2). Thereby, $\boldsymbol{f}' = \boldsymbol{f}$.*

Conversely, let us assume that $\boldsymbol{B}_\alpha(\mathcal{S}) \cap \mathcal{F}^N = \{\boldsymbol{f}\}$. If $\boldsymbol{f}' \in \mathcal{F}^N$ satisfies $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{f}'$, then $\boldsymbol{f}' \in \boldsymbol{B}_\alpha(\mathcal{S})$ by Theorem III.2 (i), which implies that $\boldsymbol{f}' = \boldsymbol{f}$.

***Proof of statement* (ii):** *Suppose that $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{f}$ for $\boldsymbol{f}$ randomly chosen with uniform distribution in $\mathcal{F}^N$. According to statement (i), the probability that $\boldsymbol{B}_\alpha(\mathcal{S}) \cap \mathcal{F}^N = \{\boldsymbol{f}\}$ is the probability that $\boldsymbol{f}$ be the unique solution of Eq. (2). We can transform $\mathcal{F}$ into $\mathcal{G} = \left\{ \frac{\alpha_i - \alpha_1}{\alpha_p - \alpha_1} : i \in [\![1, p]\!] \right\} \subset [0, 1]$ by assuming, with no loss of generality, that $\alpha_1 < \alpha_2 < \ldots < \alpha_p$. We can thus transform $\mathcal{F}^N$ into $\mathcal{G}^N \subset [0, 1]^N$. The elements of $\mathcal{G}^N$ are all k-simple vectors. We recall that a k-simple vector in $\mathbb{R}^N$ is any element of $[0, 1]^N$ with exactly k entries in $(0, 1)$ [15, Lemma 5.2]. Via this transform, $\boldsymbol{f} \in \mathcal{F}^N$ is the unique solution to (2) if and only if its transform $\boldsymbol{f}'$ is the unique solution to $\boldsymbol{y}' = \boldsymbol{A}\boldsymbol{g}$ with $\boldsymbol{g} \in \mathcal{G}^N$, where $\boldsymbol{y}'$ is the transform of $\boldsymbol{y}$. According to [15, Theorem 1.8 and Lemma 5.2], the probability that $\boldsymbol{f}'$ be the unique k-simple solution to this transformed system is $1 - P_{N-n,N-k}$, for any $k \in [\![0, n-1]\!]$. Since $\boldsymbol{f}$ is randomly chosen with uniform distribution in $\mathcal{F}^N$, $\boldsymbol{f}'$ is randomly chosen with uniform distribution in $\mathcal{G}^N$. Therefore, the probability that $\boldsymbol{f}'$ be k-simple is $\binom{N}{k} \left(\frac{2}{p}\right)^{N-k} \left(\frac{p-2}{p}\right)^{k}$. By Bayes's axiom, the probability that $\boldsymbol{f}'$ is k-simple and the unique solution to Eq. (2) is $\binom{N}{k} \left(\frac{2}{p}\right)^{N-k} \left(\frac{p-2}{p}\right)^{k} (1 - P_{N-n,N-k})$. Thereby, Eq. (17) is actually the probability that $\boldsymbol{f}$ be the unique solution to Eq. (2).*

***Proof of statement* (iii):** *We have $Q_{n,N}(p) = \mathbb{E}\big[1 -$*

$P_{N-n,N-K_N}\big]$, *where* K_N *follows the binomial distribution with parameters* N *and* $\frac{p-2}{p}$ *with mean* $N\frac{p-2}{p}$ *and variance* $N\frac{2(p-2)}{p^2}$. *Moreover, we remark that* $1-P_{N-n,N-k}=P_{n-k,N-k}$ *for all* $k\in[\![0,n-1]\!]$. *Therefore,* $Q_{n,N}(p)=\mathbb{E}\big[P_{n-K_N,N-K_N}\big]$.

For $p=2$, *Eq. (17) reduces to* $Q_{n,N}(2)=P_{n,N}$ *and the result straightforwardly follows from [8]. Henceforth, we assume* $p>2$.

By using Hoeffding's inequality [17, Theorem 1, Eq.(2.2)] with $\mu=1/2$, $g(\mu)=2$ *and* $t=\frac{2n-N-1}{2(N-1)}$, *we have:*

$$\left\{\begin{array}{lcll} P_{n,N} & \geqslant & 1-e_{n,N} & \text{for} \quad 2n>N+1 \\ P_{n,N} & \leqslant & e_{n,N} & \text{for} \quad 2n<N+1 \end{array}\right.$$

with $e_{n,N}=\exp\left(-\frac{1}{2}\frac{(2n-N-1)^2}{N-1}\right)$. *It follows that:*

$$\left\{\begin{array}{lcll} P_{n-k,N-k} & \geqslant & 1-e_{n-k,N-k} & \text{for} \quad k<2n-N-1 \\ P_{n-k,N-k} & \leqslant & e_{n-k,N-k} & \text{for} \quad k>2n-N-1 \end{array}\right.$$

First, suppose that $\frac{n}{N}<\frac{1}{2}$. *We then have* $P_{n-K_N,N-K_N}\leqslant e_{n-K_N,N-K_N}$, *which implies that* $Q_{n,N}(p)\leqslant\mathbb{E}\left[e_{n-K_N,N-K_N}\right]$. *Moreover,* $e_{n-k,N-k}\leqslant e_{n,N}$ *for any* $k\in[\![0,n-1]\!]$, *so that* $Q_{n,N}(p)\leqslant e_{n,N}$. *Since* $e_{n,N}\to 0$, *we conclude that* $\lim_{N\to+\infty}Q_{n,N}(p)=0$ *when* $\frac{n}{N}<\frac{1}{2}$ *and* (n,N) *grows proportionally.*

We now suppose that $\frac{1}{2}<\frac{n}{N}<\frac{p-1}{p}$. *We can write that:*

$$Q_{n,N}(p)=\mathbb{E}\left[P_{n-K_N,N-K_N}\mathbb{1}_{[0\leqslant K_N\leqslant 2n-N-1]}\right]+\mathbb{E}\left[P_{n-K_N,N-K_N}\mathbb{1}_{[2n-N\leqslant K_N\leqslant n-1]}\right]. \quad (18)$$

On the one hand, $\mathbb{E}\left[P_{n-K_N,N-K_N}\mathbb{1}_{[0\leqslant K_N\leqslant 2n-N-1]}\right]\leqslant\mathbb{P}\left[K_N\leqslant 2n-N-1\right]$. *On the other hand, there exists* $0<\varepsilon<\frac{p-2}{2p}$ *such that* $\frac{1}{2}<\frac{n}{N}<\frac{p-1}{p}-\varepsilon$ *so that:*

$$\mathbb{P}\big[K_N\leqslant 2n-N-1\big]\leqslant\mathbb{P}\left[K_N\leqslant N\left(\frac{p-2}{p}-2\varepsilon\right)-1\right].$$

Since $\frac{1}{N}K_N$ *is asymptotically normal [18, Sec. 1.5.5] — and we write that* $\frac{1}{N}K_N\sim\mathcal{AN}\left(\frac{p-2}{p},\frac{2(p-2)}{Np^2}\right)$ —, *it follows that for any* $\eta>0$:

$$\left|\mathbb{P}\left[K_N\leqslant N\left(\frac{p-2}{p}-2\varepsilon-\frac{1}{N}\right)\right]-\mathbb{F}_{\mathcal{N}\left(\frac{p-2}{p},\frac{2(p-2)}{Np^2}\right)}\left(\frac{p-2}{p}-2\varepsilon-\frac{1}{N}\right)\right|\leqslant\eta.$$

for N *large enough, where* $\mathbb{F}_{\mathcal{N}\left(\frac{p-2}{p},\frac{2(p-2)}{Np^2}\right)}$ *stands for the cumulative distribution function (cdf) of the normal distribution* $\mathcal{N}\left(\frac{p-2}{p},\frac{2(p-2)}{Np^2}\right)$. *With the same notation as in [18, Sec. 1.5.4], the convergence in distribution:*

$$\mathcal{N}\left(\frac{p-2}{p},\frac{2(p-2)}{Np^2}\right)\xrightarrow{d}\mathbb{1}_{[\frac{p-2}{p},\infty[}, \quad (19)$$

when N *tends to* ∞ *and Slutsky's theorem [18, Sec. 1.5.4, p. 19] imply that:*

$$\lim_{N\to\infty}\mathbb{F}_{\mathcal{N}\left(\frac{p-2}{p},\frac{2(p-2)}{Np^2}\right)}\left(\frac{p-2}{p}-2\varepsilon-\frac{1}{N}\right)=\mathbb{1}_{[\frac{p-2}{p},\infty[}\left(\frac{p-2}{p}-2\varepsilon\right)=0$$

Thence, $\lim_{N\to\infty}\mathbb{E}\left[P_{n-K_N,N-K_N}\mathbb{1}_{[0\leqslant K_N\leqslant 2n-N-1]}\right]=0$ *since* η *is arbitrary. Now, the* 2^{nd} *term in Eq. (18) tends to* 0 *as well when* N *tends to* ∞ *since* $\mathbb{E}\left[P_{n-K_N,N-K_N}\mathbb{1}_{[2n-N\leqslant K_N\leqslant n-1]}\right]\leqslant\mathbb{E}\left[e_{n-K_N,N-K_N}\right]$ *and* $e_{n-k,N-k}\leqslant\exp(-\frac{1}{2(N-1)})$ *for* $2n-N\leqslant k\leqslant n-1$.

Let us now consider the case $\frac{p-1}{p}<\frac{n}{N}<1$. *Similarly to the foregoing, we can write that:*

$$\mathbb{E}\big[P_{n-K_N,N-K_N}\big]=\mathbb{E}\left[P_{n-K_N,N-K_N}\mathbb{1}_{[0\leqslant K_N\leqslant 2n-N-2]}\right]+\mathbb{E}\left[P_{n-K_N,N-K_N}\mathbb{1}_{[2n-N-1\leqslant K_N\leqslant n-1]}\right] \quad (20)$$

First, we have $P_{n-K_N,N-K_N}\mathbb{1}_{[0\leqslant K_N\leqslant 2n-N-2]}\geqslant(1-e_{n-K_N,N-K_N})\mathbb{1}_{[0\leqslant K_N\leqslant 2n-N-2]}$. *Since we have* $1-e_{n-k,N-k}\geqslant 1-\exp(-\frac{1}{2(N-1)})$ *again, it follows that:*

$$P_{n-K_N,N-K_N}\mathbb{1}_{[0\leqslant K_N\leqslant 2n-N-2]}\geqslant\left(1-\exp\left(-\frac{1}{2(N-1)}\right)\right)\mathbb{1}_{[0\leqslant K_N\leqslant 2n-N-2]}$$

and

$$\mathbb{E}\left[P_{n-K_N,N-K_N}\mathbb{1}_{[0\leqslant K_N\leqslant 2n-N-2]}\right]\geqslant\left(1-\exp\left(-\frac{1}{2(N-1)}\right)\right)\mathbb{P}\big[K_N\leqslant 2n-N-2\big]. \quad (21)$$

There exists $0<\varepsilon<1/p$ *such that* $\frac{p-1}{p}<\frac{p-1}{p}+\varepsilon<\frac{n}{N}<1$. *Thereby,*

$$\mathbb{P}\big[K_N\leqslant 2n-N-2\big]\geqslant\mathbb{P}\big[K_N\leqslant\left(\frac{p-2}{p}+2\varepsilon\right)N-2\big].$$

Given $\eta>0$, *the asymptotic normality of* $\frac{1}{N}K_N$ *implies that:*

$$\left|\mathbb{P}\left[K_N\leqslant N\left(\frac{p-2}{p}+2\varepsilon-\frac{2}{N}\right)\right]-\mathbb{F}_{\mathcal{N}\left(\frac{p-2}{p},\frac{2(p-2)}{Np^2}\right)}\left(\frac{p-2}{p}+2\varepsilon-\frac{2}{N}\right)\right|\leqslant\eta$$

for N *large enough. Slutsky's theorem [18, Sec. 1.5.4, p. 19] and the weak convergence (19) induce that:*

$$\lim_{N\to\infty}\mathbb{F}_{\mathcal{N}\left(\frac{p-2}{p},\frac{2(p-2)}{Np^2}\right)}\left(\frac{p-2}{p}+2\varepsilon-\frac{2}{N}\right)=\mathbb{1}_{[\frac{p-2}{p},\infty[}\left(\frac{p-2}{p}+2\varepsilon\right)=1.$$

Since η *is arbitrary, we have* $\lim_{N\to\infty}\mathbb{P}\left[K_N\leqslant N\left(\frac{p-2}{p}+2\varepsilon-\frac{2}{N}\right)\right]=1$. *By injecting this result into (21), we obtain that* $\lim_{N\to\infty}\mathbb{E}\left[P_{n-K_N,N-K_N}\mathbb{1}_{[0\leqslant K_N\leqslant 2n-N-2]}\right]=1$. *It follows that the* 2^{nd} *term in Eq. (20) tends to* 0 *and the proof is complete.*

IV. Complexity analysis

In practice, the minimization of each cost considered above requires linear programming. A well-known and typical toolbox such as CVX [19], [20] relies on the interior point method, whose complexity for problems such as those treated in this paper is given by [21]. Specifically, a convex optimization

problem over $\mathbb{R}^m$ under d constraints requires, in the worst case, $\mathcal{O}(\sqrt{d})$ iterations for a computational cost of order $\mathcal{O}(m^2 d)$ per iteration and, thus, to a total computational cost of order $\mathcal{O}(m^2 d^{3/2})$. Applied to the 3 convex optimization problems treated in the paper, we obtain the following computational costs:

According to these estimations, (P_∞) and $(P_{\mathcal{F},1})$ have roughly the same computational cost, although the former is theoretically more costly than the later without any programming optimization. On the other hand, $(TP_{\mathcal{F},1})$ becomes significantly more costly than the other two, when p increases. The experimental results of the next section, beyond reconstruction assessment, provides computational time estimations on a standard computer.

V. Experimental results

The following simulation results are aimed at illustrating the theoretical framework exposed above and make it possible to assess the regularization- and transform-based approaches in terms of complexity and performance for different alphabet sizes. This assessment will also involve Mangasarian & Recht's approach [8] as a reference in the specific case $p = 2$. Thanks to these comparisons, we will conclude on the use and relevance of these different methods.

The experimental set-up, common to all the simulations whose results are given below, is the following one. We use even values for p and choose $\mathcal{F} = \{\pm(2k-1) : k = 1, \ldots, p/2\}$. For each simulation, we fix $N \in \{128, 256, 512\}$ and make n vary so as to assess a significant number of values for ratio n/N. For each pair (n, N) and each method assessed, 1000 iterations of the experiment are carried out. For each iteration, we generate a realization of the generic random matrix $\boldsymbol{A}$ with size $n \times N$ by drawing its entries from the normal distribution. We then generate a vector $\boldsymbol{f}$ with entries drawn uniformly from $\mathcal{F}$. Once $\boldsymbol{f}$ is generated, we compute $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{f}$ and solve (P_∞) — when $p = 2$ —, $(P_{\mathcal{F},1})$ and $(TP_{\mathcal{F},1})$. We compute the solutions to these optimization problems by using the Matlab CVX toolbox [19], [20], a package for solving convex problems. Finally, we compare the solution $\hat{\boldsymbol{f}}$ returned for a given optimization problem to the true signal $\boldsymbol{f}$. The recovery is said to be correct if the relative error $\frac{\|\hat{\boldsymbol{f}}-\boldsymbol{f}\|_2}{\|\boldsymbol{f}\|_2}$ is less than 10^{-6}.

Figures 1, 2 and 3 are the phase diagrams for the case $p = 2$. They involve Mangasarian and Recht's approach [8]. These phase diagrams show that the three methods perform equivalently. They also corroborate Theorems III.1 and III.3. In particular, we observe that the breakpoint is actually $n/N = 1/2$, as established theoretically. In accordance with Table I, the optimization problem $(P_{\mathcal{F},1})$ is computationally less costly than the other two. This is illustrated by Figures 4, 5 and 6. These results were obtained by using a PC with OS Linux Ubuntu 14.04 with processor Intel Core i3-2350M 2.3 GHz and 8 GB of RAM memory. The values of ratio n/N considered in these figures are those for which the algorithms under consideration recover the solution with high probability, in accordance with the phase diagrams and the theoretical results of Theorems III.1 and III.3. According to the foregoing, it is recommended to use the regularization-based approach for $p = 2$, since this method provides the least computational load for the recovery performance guaranteed by the theoretical results.

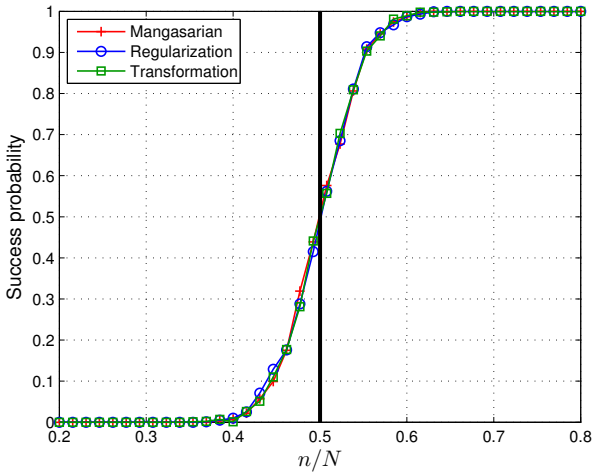

Figure 1. Phase diagrams of Mangasarian and Recht's approach [8], regularization- and transform-based methods for the case $p = 2$ and $N = 128$.

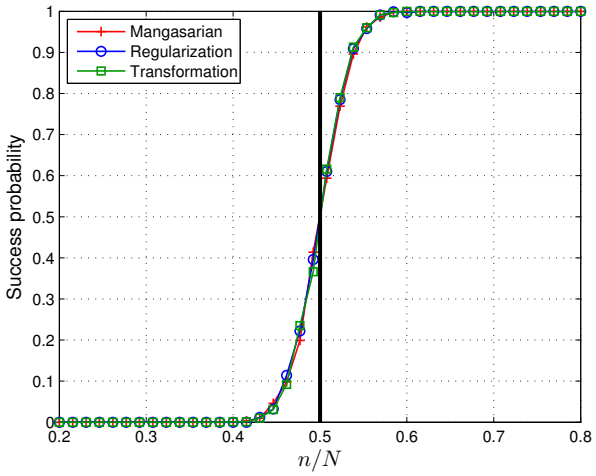

Figure 2. Phase diagrams of Mangasarian and Recht's approach [8], regularization- and transform-based methods for the case $p = 2$ and $N = 256$.

We now address the case $p > 2$, for which the approach in [8] is not applicable. Figures 7, 8, 9 and 10 provide the phase diagrams of the regularization-based and transform-based methods for $p = 4$, $p = 6$ and different values of N. The regularization-based approach has a much higher breakpoint and thus performs less well than the transform-based one. The breakpoint of the transform-based approach is actually the value given by Theorem III.3. On the other hand, when p increases, the transform-based approach deals with much higher dimensions than the regularization-based one. Therefore, the computational time required by the former

Table I
COMPUTATIONAL COST ANALYSIS

	Dimension	# constraints	# iterations	Computational cost per iteration	Total
(P_∞)	N	$2N+n$	$\mathcal{O}(\sqrt{2N+n})$	$\mathcal{O}(N^2(2N+n))$	$\mathcal{O}(N^2(2N+n)^{3/2})$
$(P_{\mathcal{F},1})$	N	n	$\mathcal{O}(\sqrt{n})$	$\mathcal{O}(N^2 n)$	$\mathcal{O}(N^2 n^{3/2})$
$(TP_{\mathcal{F},1})$	pN	$N+n$	$\mathcal{O}(\sqrt{N+n})$	$\mathcal{O}(p^2N^2(N+n))$	$\mathcal{O}(p^2N^2(N+n)^{3/2})$

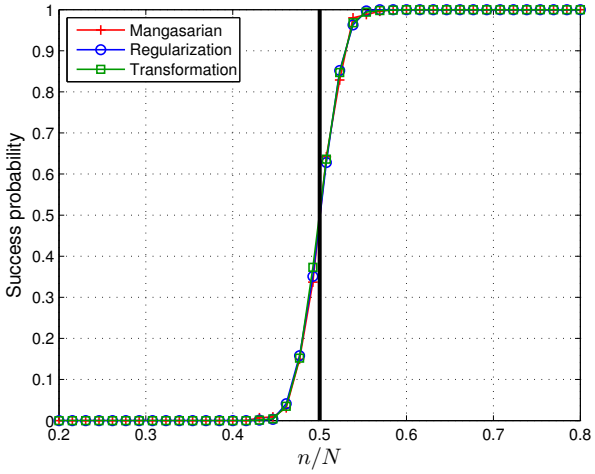

Figure 3. Phase diagrams of Mangasarian and Recht's approach [8], regularization- and transform-based methods for the case $p = 2$ and $N = 512$.

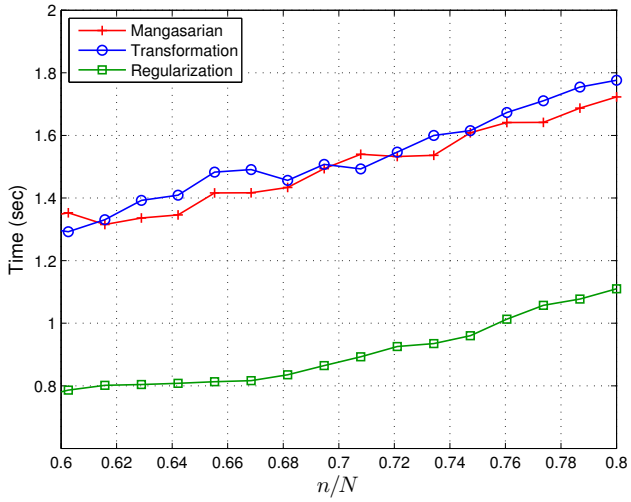

Figure 5. Computational time in seconds of Mangasarian and Recht's approach [8], regularization- and transform-based methods for the case $p = 2$ and $N = 256$.

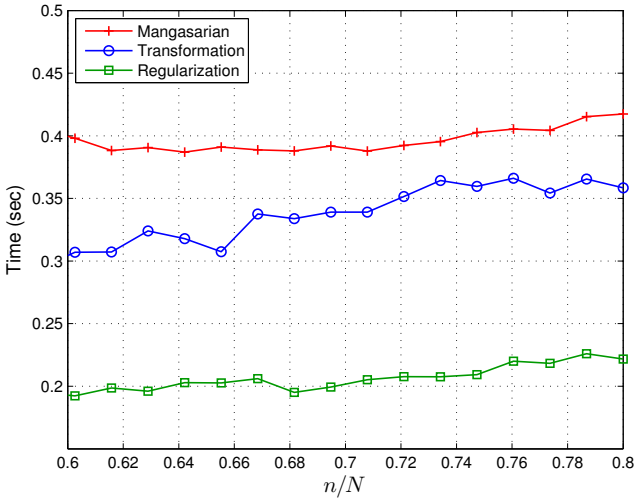

Figure 4. Computational time in seconds of Mangasarian and Recht's approach [8], regularization- and transform-based methods for the case $p = 2$ and $N = 128$.

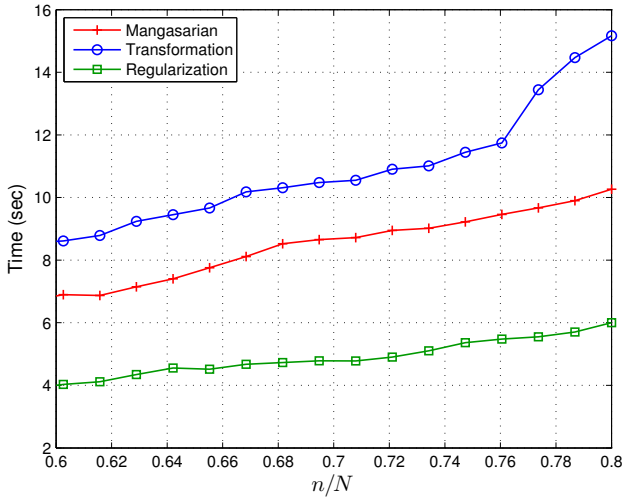

Figure 6. Computational time in seconds of Mangasarian and Recht's approach [8], regularization- and transform-based methods for the case $p = 2$ and $N = 512$.

can significantly be larger than that of the latter, as predicted by the complexity analysis of Section IV.

VI. DISCUSSION

The ℓ_∞-norm optimization of [8], the regularization-based

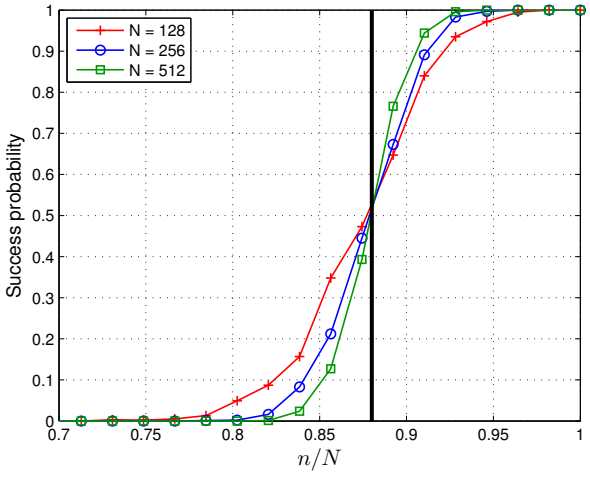

Figure 7. Phase diagrams of the regularization-based method, for $p = 4$ and $N \in \{128, 256, 512\}$.

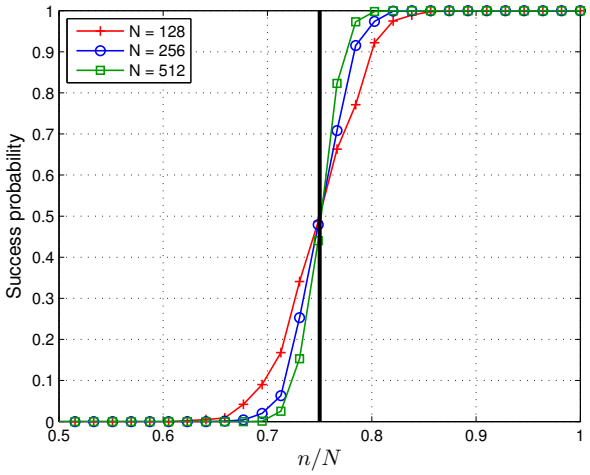

Figure 8. Phase diagrams of the transform-based method, for $p = 4$ and $N \in \{128, 256, 512\}$.

method of Section III-A and the transforma-tion-based method of Section III-B involve minimizing (P_∞), $(P_{\mathcal{F},1})$ and $(TP_{\mathcal{F},1})$, respectively. The transform-based method requires an additional step, which is the application of $\boldsymbol{B}_\alpha$ to the set $\mathcal{S}$ of solutions to $(TP_{\mathcal{F},1})$. In this discussion, as well as the conclusion of the paper, we designate each of these three methods by the criterion it minimizes. This slight language abuse is convenient without entailing any confusion.

A. Performance and complexity assessment

For $p = 2$, Theorems III.1 and III.3, corroborated by experimental results, guarantee that $(P_{\mathcal{F},1})$ and $(TP_{\mathcal{F},1})$ yield same recovery performance as (P_∞). Moreover, the complexity analysis of Section IV and the experimental computational times show that $(P_{\mathcal{F},1})$ is slightly less computationally expensive than $(TP_{\mathcal{F},1})$ and (P_∞). Therefore, for $p = 2$, $(P_{\mathcal{F},1})$ should be used instead of the other two.

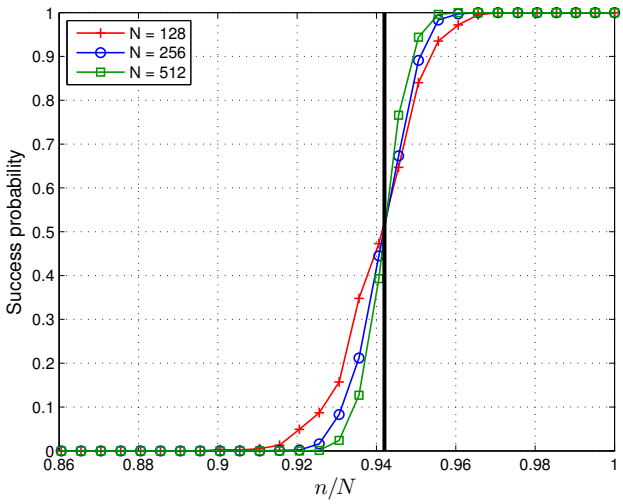

Figure 9. Phase diagrams of the regularization-based method, for $p = 6$ and $N \in \{128, 256, 512\}$.

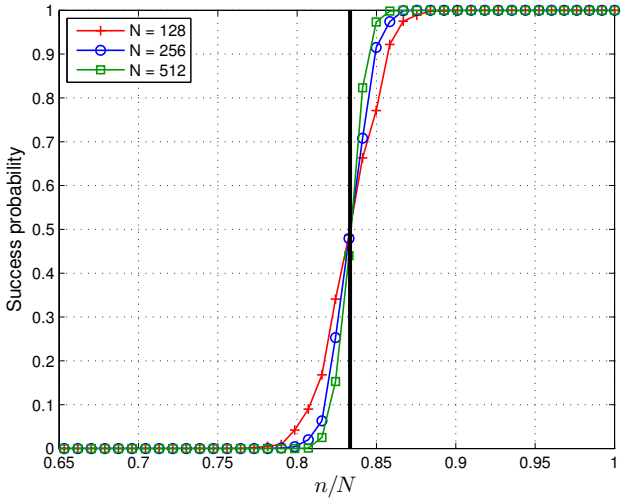

Figure 10. Phase diagrams of the transform-based method, for $p = 6$ and $N \in \{128, 256, 512\}$.

For $p > 2$, (P_∞) does not apply. In addition, the recovery performance of $(TP_{\mathcal{F},1})$ is experimentally better than that of $(P_{\mathcal{F},1})$ and the complexity analysis shows that the former is less costly than the latter. Therefore, the choice between the two methods proposed in this paper depends on the application: if more emphasis must be given to recovery performance, then $(TP_{\mathcal{F},1})$ is better than $(P_{\mathcal{F},1})$ and if more emphasis on computational load is required, $(P_{\mathcal{F},1})$ can be preferred.

B. Complementary remarks

Via the following three remarks, we now discuss to what extent some notions and results, complementary to those mentioned above, relate to the theoretical framework developed in the present paper.

1) We recall that the Kruskal rank $\mathrm{Krank}(\boldsymbol{A})$ of any given matrix $\boldsymbol{A}$ is the maximal L such that every L columns are linearly independent. The computation of $\mathrm{Krank}(\boldsymbol{A})$ is NP-complete [22]. On the other hand, it is known that the system $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$ of linear equations is well-posed for k-sparse vectors $\boldsymbol{x}$ if and only if $2k \leq \mathrm{Krank}(\boldsymbol{A})$. For the transform-based model of Eq. (10), we derive from Theorem B.1 that $\mathrm{Krank}(\boldsymbol{\Phi}) \geq 2N$ since $\boldsymbol{s}(\boldsymbol{f})$ is N-sparse.

2) Since $\boldsymbol{s}(\boldsymbol{f})$ belongs to $\{0,1\}^{Np}$, $\boldsymbol{s}(\boldsymbol{f})$ is 0-simple in the sense given in [15, Section 5.2, p. 540]. By applying [15, Theorem 1.8 & Lemma 5.2], the probability that $\boldsymbol{s}(\boldsymbol{f})$ be the unique solution to $\boldsymbol{A}\boldsymbol{B}_\alpha\boldsymbol{u} = \boldsymbol{y}$ is 0. Therefore, the constraint $\boldsymbol{B}_1\boldsymbol{u} = \boldsymbol{1}_N$ plays a crucial role in the optimization problems $(TP_{\mathcal{F},0})$ and $(TP_{\mathcal{F},1})$ to guarantee the uniqueness of the solution in Proposition III.2 and Theorem III.3.

3) Although RIP conditions [23], [24] are the most privileged sufficient conditions to guarantee the uniqueness of a sparse solution to ℓ_1-norm optimization problems, these conditions are not appropriate to study the relaxation from $(TP_{\mathcal{F},0})$ to $(TP_{\mathcal{F},1})$. Indeed, for $p = 2$, Theorem B.1 states that the N-sparse vector $\boldsymbol{s}(\boldsymbol{f})$ is the unique solution to both $(TP_{\mathcal{F},0})$ and $(TP_{\mathcal{F},1})$ in $\mathbb{R}^{2N}$. For $p > 2$, the two problems are not equivalent because experiments show that the set of solutions to $(TP_{\mathcal{F},1})$ do not involve $\boldsymbol{s}(\boldsymbol{f})$ only.

VII. Conclusion

Two frameworks have been proposed for the underdetermined source separation problem of finite alphabet signals. The first one is based on regularization and the second one relies on a suitable sparse transform. Both frameworks are based on convex relaxation aimed at recovering the ideal finite alphabet signal by solving ℓ_1-norm optimization problems. Simulation results illustrate the effectiveness of the proposed approaches, although the RIP condition is not satisfied. For $p = 2$, the regularization-based approach should be used for reasons detailed in Section VI. For $p > 2$, the computational cost of the regularization-based approach remains lesser than that of the transform-based one. However, the recovery performance measurements of the latter exceed those of the former.

As mentioned in the introduction, the results presented above apply to various problems such as source separation, wireless communication systems, operations research. In particular, they can help choose the number of transmitters and receivers, as well as the type of modulation, in wireless communication systems, even in presence of noise. For instance, [25], [26] deals with signal recovery in massive MIMO systems, where finite alphabets are used and the decoding can be performed by underdetermined sparse source recovery via ℓ_1-norm minimization. In [25], [26], noise is taken into account by slightly modifying the criterion of the transform-based approach. As a continuation of the results exposed in [25], [26], a full theoretical study dedicated to the extension of the present work to noisy signals is in-progress. In this respect, a comparison with the approach proposed in [27] should be undertaken. Then, a combination of the two approaches could be particularly relevant in a communication context. Indeed, prior knowledge of the communication signals can be taken into account so as to define new transforms beyond the full disjunctive coding exploited in the present paper.

Acknowledgement

The authors are very grateful to the reviewers and the Associate Editor, whose comments and suggestions helped us improve significantly this paper.

Appendix A
Proof of equality (16)

A. Proof of direct inclusion $\mathrm{Ker}\,\boldsymbol{\Phi} \cap \mathcal{K}_{\boldsymbol{f}}^{N,p} \subset \mathrm{Ker}\,\boldsymbol{\Phi} \cap \mathcal{C}_{\boldsymbol{f}}^{N,p}$.

For $j \in \{1, 2, \ldots, Np\}$, let $s_j(\boldsymbol{f})$ be the j^{th} coordinate of $\boldsymbol{s}(\boldsymbol{f})$. Given any $\boldsymbol{h} \in \mathcal{K}_{\boldsymbol{f}}^{N,p}$, we have

$$\begin{aligned}\|\boldsymbol{s}(\boldsymbol{f}) + \boldsymbol{h}\|_1 &= \sum_{i=1}^{N} \sum_{j \in T_i} |s_j(\boldsymbol{f}) + h_j| \\ &= \sum_{i=1}^{N} \left(\sum_{j \in T_i \setminus \{n_i(\boldsymbol{f})\}} |h_j| + \left|h_{n_i(\boldsymbol{f})} + 1\right| \right)\end{aligned}$$

since $s_j(\boldsymbol{f}) = 0$ if $j \neq n_i(\boldsymbol{f})$ and $s_j(\boldsymbol{f}) = 1$, otherwise. According to the definition of $\mathcal{K}_{\boldsymbol{f}}^{N,p}$, the absolute values are inconsequential in the last equality above. Thereby,

$$\begin{aligned}\|\boldsymbol{s}(\boldsymbol{f}) + \boldsymbol{h}\|_1 &= \sum_{i=1}^{N} \left(\sum_{j \in T_i \setminus \{n_i(\boldsymbol{f})\}} h_j + h_{n_i(\boldsymbol{f})} + 1 \right) \\ &= N + \sum_{i=1}^{N} \sum_{j \in T_i} h_j = N\end{aligned}$$

by definition of $\mathcal{K}_{\boldsymbol{f}}^{N,p}$. Therefore, $\mathcal{K}_{\boldsymbol{f}}^{N,p} \subset \mathcal{C}_{\boldsymbol{f}}^{N,p}$, which implies that

$$\mathrm{Ker}\,\boldsymbol{\Phi} \cap \mathcal{K}_{\boldsymbol{f}}^{N,p} \subset \mathrm{Ker}\,\boldsymbol{\Phi} \cap \mathcal{C}_{\boldsymbol{f}}^{N,p} \tag{22}$$

B. Proof of converse inclusion $\mathrm{Ker}\,\boldsymbol{\Phi} \cap \mathcal{C}_{\boldsymbol{f}}^{N,p} \subset \mathrm{Ker}\,\boldsymbol{\Phi} \cap \mathcal{K}_{\boldsymbol{f}}^{N,p}$.

Let $\boldsymbol{h}$ be any element of $\mathrm{Ker}\,\boldsymbol{B}_1 \cap \mathcal{C}_{\boldsymbol{f}}^{N,p}$. We have:

$$\|\boldsymbol{s}(\boldsymbol{f}) + \boldsymbol{h}\|_1 = N \quad \text{and} \quad \boldsymbol{B}_1\boldsymbol{h} = \boldsymbol{0} \tag{23}$$

We derive from Eq. (9) that

$$\boldsymbol{B}_1(\boldsymbol{s}(\boldsymbol{f}) + \boldsymbol{h}) = \boldsymbol{1}_N. \tag{24}$$

According to Eqs. (23), (24) and Lemma A.1, $\boldsymbol{h} \in \mathcal{K}_{\boldsymbol{f}}^{N,p}$. We thus have $\mathrm{Ker}\,\boldsymbol{B}_1 \cap \mathcal{C}_{\boldsymbol{f}}^{N,p} \subset \mathcal{K}_{\boldsymbol{f}}^{N,p}$. Since $\mathrm{Ker}\,\boldsymbol{\Phi} \subset \mathrm{Ker}\,\boldsymbol{B}_1$, it follows from the foregoing that

$$\mathrm{Ker}\,\boldsymbol{\Phi} \cap \mathcal{C}_{\boldsymbol{f}}^{N,p} = \mathrm{Ker}\,\boldsymbol{\Phi} \cap \mathrm{Ker}\,\boldsymbol{B}_1 \cap \mathcal{C}_{\boldsymbol{f}}^{N,p} \subset \mathrm{Ker}\,\boldsymbol{\Phi} \cap \mathcal{K}_{\boldsymbol{f}}^{N,p}.$$

Lemma A.1 *The coordinates of any $\boldsymbol{u} \in \mathbb{R}^{Np}$ such that $\boldsymbol{B}_1\boldsymbol{u} = \boldsymbol{1}_N$ and $\|\boldsymbol{u}\|_1 = N$ are non-negative.*

Proof *If $\boldsymbol{u} = [u_1, u_2, \ldots, u_{Np}] \in \mathbb{R}^{Np}$ is such that $\boldsymbol{B}_1\boldsymbol{u} = \boldsymbol{1}_N$, then $\sum_{j \in T_i} u_j = 1$ for every $i \in \{1, 2, \ldots, N\}$. It follows that $\sum_{i=1}^{N} \sum_{j \in T_i} u_j = N$. If, in addition, $\|\boldsymbol{u}\|_1 = N$, we have $\sum_{i=1}^{N} \sum_{j \in T_i} |u_j| = \sum_{i=1}^{N} \sum_{j \in T_i} u_j$, which implies that each $u_j \geq 0$.*

APPENDIX B

Theorem B.1 *For $\mathcal{F} = \{\alpha_1, \alpha_2\}$ ($p = 2$) where $\alpha_1 < \alpha_2$, $\boldsymbol{f} \in \mathcal{F}^N$ is the unique solution to Eq. (2) if and only if $\boldsymbol{s}(\boldsymbol{f})$ is the unique solution to $(TP_{\mathcal{F},1})$.*

Proof *Given any $\boldsymbol{f} \in \mathcal{F}^N$ and any $i \in [\![1, N]\!]$, we define:*

$$\varepsilon_i(\boldsymbol{f}) = \begin{cases} 1, & \text{if} \quad n_i(\boldsymbol{f}) = 2i - 1 \\ -1, & \text{if} \quad n_i(\boldsymbol{f}) = 2i \end{cases}.$$

Let $\boldsymbol{h}$ be any element of $\mathcal{K}_{\boldsymbol{f}}^{N,2}$. Since $T_i = \{2i-1,\, 2i\}$ for any $i \in [\![1, N]\!]$, we set

$$\gamma_i = |h_{2i-1}| = |h_{2i}| \in [0, 1].$$

We have:

$$h_{2i-1} = -\varepsilon_i(\boldsymbol{f})\, \gamma_i \quad \text{and} \quad h_{2i} = \varepsilon_i(\boldsymbol{f})\, \gamma_i.$$

Then $\boldsymbol{h} = (\boldsymbol{D}_{\varepsilon(\boldsymbol{f})}\boldsymbol{\gamma}) \otimes [-1 \;\; 1]^T$ with $\boldsymbol{\gamma} = (\gamma_1, \cdots, \gamma_N)^T$ and

$$\boldsymbol{D}_{\varepsilon(\boldsymbol{f})} = \operatorname{diag}(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_N) \tag{25}$$

Conversely, any vector $(\boldsymbol{D}_{\varepsilon(\boldsymbol{f})}\boldsymbol{\gamma}) \otimes [-1 \;\; 1]^T$ where $\boldsymbol{\gamma} \in [0, 1]^N$ belongs to $\mathcal{K}_{\boldsymbol{f}}^{N,2}$. Therefore,

$$\mathcal{K}_{\boldsymbol{f}}^{N,2} = \left\{ (\boldsymbol{D}_{\varepsilon(\boldsymbol{f})}\boldsymbol{\gamma}) \otimes [-1 \;\; 1]^T : \; \boldsymbol{\gamma} \in [0, 1]^N \right\}. \tag{26}$$

We now prove that (12) when $p = 2$ is equivalent to $\mathcal{K}_{\boldsymbol{f}}^{N,2} \cap \operatorname{Ker}(\boldsymbol{A}\boldsymbol{B}_\alpha) = \{\boldsymbol{0}\}$. In other words, we want to prove the equivalence:

$$\left[\mathcal{K}_{\boldsymbol{f}}^{N,2} \cap \operatorname{Ker} \boldsymbol{\Phi} \;\subset\; \operatorname{Ker} \boldsymbol{B}_\alpha\right] \Longleftrightarrow \left[\mathcal{K}_{\boldsymbol{f}}^{N,2} \cap \operatorname{Ker}(\boldsymbol{A}\boldsymbol{B}_\alpha) = \{\boldsymbol{0}\}\right] \tag{27}$$

In the sequel, we use the following property. For any $\boldsymbol{h} = (\boldsymbol{D}_{\varepsilon(\boldsymbol{f})}\boldsymbol{\gamma}) \otimes [-1 \;\; 1]^T \in \mathcal{K}_{\boldsymbol{f}}^{N,2}$, it follows from (8) and usual properties of the Kronecker product that:

$$\begin{aligned} \boldsymbol{B}_\alpha \boldsymbol{h} &= \left(\boldsymbol{I}_N \otimes \boldsymbol{\alpha}^T\right)\left((\boldsymbol{D}_{\varepsilon(\boldsymbol{f})}\boldsymbol{\gamma}) \otimes [-1 \;\; 1]^T\right) \\ &= \left(\boldsymbol{I}_N\, \boldsymbol{D}_{\varepsilon(\boldsymbol{f})}\, \boldsymbol{\gamma}\right) \otimes \left(\boldsymbol{\alpha}^T [-1 \;\; 1]^T\right) \\ &= (\alpha_2 - \alpha_1)\, \boldsymbol{D}_{\varepsilon(\boldsymbol{f})}\, \boldsymbol{\gamma}. \end{aligned} \tag{28}$$

We begin by the direct implication in (27). So, we suppose that $\mathcal{K}_{\boldsymbol{f}}^{N,2} \cap \operatorname{Ker} \boldsymbol{\Phi} \subset \operatorname{Ker} \boldsymbol{B}_\alpha$. By definition of $\boldsymbol{\Phi}$, $\operatorname{Ker} \boldsymbol{\Phi} = \operatorname{Ker} \boldsymbol{B}_1 \cap \operatorname{Ker}(\boldsymbol{A}\boldsymbol{B}_\alpha)$ and by definition [1] of $\mathcal{K}_{\boldsymbol{f}}^{N,2}$, $\mathcal{K}_{\boldsymbol{f}}^{N,2} \subset \operatorname{Ker} \boldsymbol{B}_1$. Therefore:

$$\begin{aligned} \mathcal{K}_{\boldsymbol{f}}^{N,2} \cap \operatorname{Ker} \boldsymbol{\Phi} &= \mathcal{K}_{\boldsymbol{f}}^{N,2} \cap \operatorname{Ker} \boldsymbol{B}_1 \cap \operatorname{Ker}(\boldsymbol{A}\boldsymbol{B}_\alpha) \\ &= \mathcal{K}_{\boldsymbol{f}}^{N,2} \cap \operatorname{Ker}(\boldsymbol{A}\boldsymbol{B}_\alpha) \end{aligned} \tag{29}$$

Since $\mathcal{K}_{\boldsymbol{f}}^{N,2} \cap \operatorname{Ker}(\boldsymbol{A}\boldsymbol{B}_\alpha)$ in not empty because the null vector $\boldsymbol{0}$ of $\mathbb{R}^N$ belongs to it, let $\boldsymbol{h}$ be one of its elements. We then have $\boldsymbol{h} \in \mathcal{K}_{\boldsymbol{f}}^{N,2}$ and $\boldsymbol{A}\boldsymbol{B}_\alpha\boldsymbol{h} = \boldsymbol{0}$. It follows from (12) with $p = 2$ that $\boldsymbol{B}_\alpha\boldsymbol{h} = \boldsymbol{0}$. According to Eqs. (26) and (28), $\boldsymbol{h} = (\boldsymbol{D}_{\varepsilon(\boldsymbol{f})}\boldsymbol{\gamma}) \otimes [-1 \;\; 1]^T$ with $\boldsymbol{\gamma} \in [0, 1]^N$ and $\boldsymbol{B}_\alpha\boldsymbol{h} = (\alpha_2 - \alpha_1)\, \boldsymbol{D}_{\varepsilon(\boldsymbol{f})}\, \boldsymbol{\gamma}$. Since $\alpha_1 \neq \alpha_2$, $\boldsymbol{B}_\alpha\boldsymbol{h} = \boldsymbol{0}$ is equivalent to $\boldsymbol{D}_{\varepsilon(\boldsymbol{f})}\, \boldsymbol{\gamma} = \boldsymbol{0}$ and the unique solution in $\boldsymbol{\gamma}$ to this equality is $\boldsymbol{\gamma} = \boldsymbol{0}$ since the determinant of $\boldsymbol{D}_{\varepsilon(\boldsymbol{f})}$ is non null. Thereby, $\boldsymbol{h} = \boldsymbol{0}$ so that $\mathcal{K}_{\boldsymbol{f}}^{N,2} \cap \operatorname{Ker}(\boldsymbol{A}\boldsymbol{B}_\alpha) = \{\boldsymbol{0}\}$. The converse is straightforward.

[1] Note that $\mathcal{K}_{\boldsymbol{f}}^{N,p} \subset \operatorname{Ker} \boldsymbol{B}_1$ for actually any p since, with the same notation as in Eq. (11), $\forall i \in [\![1, N]\!]$, $\sum_{j \in T_i} h_j = 0$

REFERENCES

[1] A. Aïssa-El-Bey, K. Abed-Meraim, and Y. Grenier, "Underdetermined blind audio source separation using modal decomposition," *EURASIP Journal on Audio, Speech, and Music Processing*, vol. 2007, no. 1, pp. 1–15, March 2007.

[2] V. Varajarajan and J.L. Krolik, "Multichannel system identification methods for sensor array calibration in uncertain multipath environments," in *IEEE Signal Processing Workshop on Statistical Signal Processing (SSP)*, Singapore, August 2001, pp. 297–300.

[3] A. Rouxel, D. Le Guennec, and O. Macchi, "Unsupervised adaptive separation of impulse signals applied to EEG analysis," in *IEEE International Conference on Acoustics, Speech, Signal Processing (ICASSP)*, Istanbul, Turkey, June 2000, vol. 1, pp. 420–423.

[4] K. Abed-Meraim, S. Attallah, T.J. Lim, and M.O. Damen, "A blind interference canceller in DS-CDMA," in *IEEE International Symposium on Spread Spectrum Techniques and Applications*, Parsippany, USA, September 2000, vol. 2, pp. 358–362.

[5] S.M. Aziz-Sbai, A. Aïssa-El-Bey, and D. Pastor, "Robust underdetermined blind audio source separation of sparse signals in the-frequency domain," in *IEEE International Conference on Acoustics, Speech, Signal Processing (ICASSP)*, Prague, Czech Republic, May 2011, pp. 3716–3719.

[6] S.M. Aziz-Sbai, A. Aïssa-El-Bey, and D. Pastor, "Contribution of statistical tests to sparseness-based blind source separation," *EURASIP journal on applied signal processing*, vol. 2012, no. 169, july 2012.

[7] Y. Li, A. Cichoki, and L. Zhang, "Blind separation and extraction of binary sources," *IEICE Transactions on Fundamentals*, vol. E86-A, no. 3, pp. 580–589, March 2003.

[8] O. L. Mangasarian and B. Recht, "Probability of unique integer solution to a system of linear equations," *European Journal of Operational Research*, vol. 214, no. 1, pp. 27–30, April 2011.

[9] Z. Tian, G. Leus, and V. Lottici, "Detection of sparse signals under finite alphabet constraints," in *IEEE International Conference on Acoustics, Speech, Signal Processing (ICASSP)*, Taipei, Taiwan, April 2009, pp. 2349–2352.

[10] T. S. Jayram, S. Pal, and V. Arya, "Recovery of a sparse integer solution to an underdetermined system of linear equations," *CoRR*, vol. abs/1112.1757, December 2011.

[11] S. M. Aziz-Sbai, A. Aïssa-El-Bey, and D. Pastor, "Underdetermined source separation of finite alphabet signals via ℓ_1 minimization," in *IEEE International Conference on Information Sciences, Signal Processing and their Applications (ISSPA)*, Montreal, Quebec, Canada, July 2012, pp. 625–628.

[12] D. L. Donoho and M. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ_1 minimization," *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, vol. 100, no. 5, pp. 2197–2202, March 2003.

[13] R. Gribonval and M. Nielson, "Sparse representations in unions of bases," *IEEE Transactions on Information Theory*, vol. 49, no. 12, pp. 3320–3325, December 2003.

[14] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Journal on Scientific Computing*, vol. 20, no. 1, pp. 33–61, 1998.

[15] D.L. Donoho and J. Taner, "Counting faces of randomly projected polytopes when the projection radically lowers dimension," *Journal of the American Mathematical Society*, vol. 22, no. 1, pp. 1–53, 2009.

[16] Hao Zhu and G.B. Giannakis, "Exploiting sparse user activity in multiuser detection," *IEEE Transactions on Communications*, vol. 59, no. 2, pp. 454–465, February 2011.

[17] W. Hoeffding, "Probability inequalities for sums of bounded random variables," *Journal of the American Statistical Association*, vol. 58, no. 301, pp. 13 – 30, March 1963.

[18] R. J. Serfling, *Approximations theorems of mathematical statistics*, Wiley, 1980.

[19] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," http://cvxr.com/cvx, March 2014.

[20] M. Grant and S. Boyd, "Graph implementations for nonsmooth convex programs," in *Recent Advances in Learning and Control*, V. Blondel, S. Boyd, and H. Kimura, Eds., Lecture Notes in Control and Information Sciences, pp. 95–110. Springer-Verlag Limited, 2008, http://stanford.edu/~boyd/graph_dcp.html.

[21] Y. Nesterov and A. Nemirovsky, *Interior Point Polynomial Algorithms in Convex Programming*, Studies in Applied Mathematics. Society for Industrial and Applied Mathematics, 1994.

[22] L.-H. Lim and P. Comon, "Multiarray signal processing: Tensor decomposition meets compressed sensing," *Comptes-Rendus de l'Académie des Sciences, Mécanique*, vol. 338, no. 6, pp. 311–320, june 2010.
[23] S. Foucart, "A note on guaranted sparse recovery via ℓ_1 minimization," *Applied and Computational Harmonic Analysis*, vol. 29, no. 1, pp. 97–103, July 2010.
[24] E. J. Candes and T. Tao, "Decoding by linear programming," *IEEE Transactions on Information Theory*, vol. 51, no. 12, pp. 4203–4215, December 2005.
[25] Y. Fadlallah, A. Aïssa-El-Bey, K. Amis, R. Pyndiah, and D. Pastor, "New decoding strategy for underdetermined MIMO transmission sparse decomposition," in *21st European Signal Processing Conference (EUSIPCO)*, September 2013.
[26] Y. Fadlallah, A. Aïssa-El-Bey, K. Amis, D. Pastor, and R. Pyndiah, "New iterative detector of MIMO transmission using sparse decomposition," *to appear in IEEE Transactions on Vehicular Technology*, 2014.
[27] X. Fan, J. Song, D. Palomar, and O. Au, "Universal binary semidefinite relaxation for ML signal detection," *IEEE Transactions on Communications*, vol. 61, no. 11, pp. 4565–4576, November 2013.

Abdeldjalil Aïssa-El-Bey (M'07, SM'12) was born in Algiers, Algeria, in 1981. He received the State Engineering degree from École Nationale Polytechnique (ENP), Algiers, Algeria, in 2003, the M.S. Degree in signal processing from Supélec and Paris XI University, Orsay, France, in 2004 and the Ph.D. degree in signal and image processing from Telecom ParisTech Paris, France in 2007. He is currently and since 2007 Associate Professor at Signal & Communications department of Telecom Bretagne. His research interests are blind source separation, blind system identification and equalization, statistical signal processing, wireless communications, and adaptive filtering.

Dominique Pastor was born in Cahors, France, in 1963. He graduated from Telecom Bretagne (Brest, France) in 1986 and from the University of Rennes (France) in 1997 (Ph.D.). From 1987 until 2000, he was with Thales. In particular, between 1990 and 1998, he was with Thales Avionics where his research concerned speech processing for applications to speech recognition systems embedded in military fast jet cockpits and, from 1998 to 2000, he was with Thales Nederland where he worked on the detection of radar targets in sea clutter. In September 2000, he joined Altran Technologies Nederland as a senior consultant. Since September 2002, he is with Institut Telecom, where he is currently Professor at Telecom Bretagne. His current research interests focus on statistical signal processing and sparse transforms with applications to physiological signals including speech.

Si Mohamed Aziz Sbaï was born in Casablanca, Morocco, in 1983. He received the State Engineering degree in telecommunication from Ecole Nationale Superieure d'Electronique, Informatique & Radiotelecommunication de Bordeaux, France, in 2007 and the Ph.D. degree in applied mathematics and information technology from Telecom Bretagne & University of Bretagne Occidentale, in 2012. From 2007 to 2009, he was a research engineer at Audionamix (a start-up company formerly known as Mist-Technologies) in Paris, working on mono & stereo source separation. Since 2012, he is with Fogale Company as an algorithm engineer, first in Nimes, France and then in Geneva, Switzerland. His research interests include inverse problem and sparsity-based signal processing.

Yasser Fadlallah (S'13) was born in Beirut, Lebanon, in 1988. He received the Telecommunication Engineering Diploma from the Lebanese University, Beirut, in 2009; the M.S. degree in signal processing from the Université de Bretagne Occidentale, Brest, France, in 2010; and the Ph.D. degree from Télécom Bretagne, Brest, in 2013. In 2012, he was visiting researcher with the Coding and Signal Transmission Laboratory, University of Waterloo, Canada. He was also R&D engineer at Orange Labs, Paris, and currently post-doctoral fellow at INRIA. His research interests are in advanced receivers for Long-Term Evolution Advanced, cross-layer optimization, interference channels, low-complexity detectors, and multiple-input-multiple-output systems.